prior to accessing the content server, executing at least one operation based on the
at least one recognized audio command.

5. (Amended) The method of claim 4 further comprising:
verifying the at least one recognized audio command.

6. (Amended) The method of claim 23 further comprising:
prior to accessing the content server, generating an error notification when the at
least one first confidence value and the at least one second confidence
values are below a minimum confidence level.

8. (Amended) The method of claim 24 further comprising:
prior to accessing a content server, generating an error notification when the at
least one terminal confidence value and the at least one network
confidence value are below a minimum confidence level.

9. (Amended) The method of claim 24 further comprising:
prior to selecting the at least one recognized audio command, weighting the at
least one terminal confidence value by a terminal weight factor and the at
least one network confidence value by a network weight factor.

10. (Amended) The method of claim 24 further comprising:
filtering the at least one recognized audio command based on the at least one
recognized audio command confidence value; and
executing an operation based on the recognized audio command having the
highest recognized audio command confidence value.

11. (Amended) The method of claim 24 further comprising:
verifying the at least one recognized audio command to generate a verified
recognized audio command; and
executing an operation based on the verified recognized audio command.

2

13. (Amended) The apparatus of claim 25 further comprising:

a the dialog manager operably coupled to the means for receiving, wherein the means for receiving selects the at least one recognized audio command having a recognized confidence value from the at least one first recognized audio command and the at least one second recognized audio command based on the at least one first confidence value and the at least one second confidence value.

15. (Amended) The apparatus of claim 25 further comprising:

wherein the dialog manager retrieves encoded information in response to the dialog manager audio command.

18. (Amended) The apparatus of claim 17 wherein when the means for receiving provides the dialog manager with an error notification, the output message is an error statement.

21. (Amended) The system of claim 26 further comprising:

wherein the dialog manager retrieves encoded information from the content server in response to the dialog manager audio command.

23. (Added 11/20/02) A method for multi-level distributed speech recognition comprising:

providing an audio command to a first speech recognition engine and at least one second speech recognition engine;

recognizing the audio command within the first speech recognition engine to generate at least one first recognized audio command, wherein the at least one first recognized audio command has a corresponding first confidence value;

3

recognizing the audio command within the at least one second speech recognition engine, independent of recognizing the audio command by the first speech recognition engine, to generate at least one second recognized audio command, wherein the at least one second recognized audio command has a corresponding second confidence value;

selecting at least one recognized audio command having a recognized audio command confidence value from the at least one first recognized audio command and the at least one second recognized audio command based on the at least one first confidence value and the at least one second confidence value; and

accessing a content server in response to the at least one recognized audio command.

24. (Added 11/20/02) A method for multi-level distributed speech recognition comprising:

providing an audio command to a terminal speech recognition engine and at least one network speech recognition engine;

recognizing the audio command within the terminal speech recognition engine to generate at least one terminal recognized audio command, wherein the at least one terminal recognized audio command has a corresponding terminal confidence value;

recognizing the audio command within the at least one network speech recognition engine to generate at least one network recognized audio command, wherein the at least one network recognized audio command has a corresponding network confidence value; and

selecting at least one recognized audio command having a recognized audio command confidence value from the at least one terminal recognized audio command and the at least one network recognized audio command; and

4

accessing a content server in response to the at least one recognized audio command.

25. (Added 11/20/02) An apparatus for multi-level distributed speech recognition comprising:

a first speech recognition means, operably coupled to an audio subsystem, for receiving an audio command and generating at least one first recognized audio command, wherein the at least one first recognized audio command has a first confidence value;

a second speech recognition means, operably coupled to the audio subsystem, for receiving the audio command and generating, independent of the first speech recognition means, at least one second recognized audio command, wherein each of the at least one second recognized audio command has a second confidence value; and

a means, operably coupled to the first speech recognition means and the second speech recognition means, for receiving the at least one first recognized audio command and the at least one second recognized audio command;

a dialog manager operably coupled to the first speech recognition means and the second speech recognition means and operably coupleable to a content server; and

the dialog manager determines a dialog manager audio command from the at least one recognized command confidence levels and wherein such that the dialog manager access the content server in response to the dialog manager audio command.

26. (Added 11/20/02) A system for multi-level distributed speech recognition comprising:

a terminal speech recognition engine operably coupled to a microphone and coupled to receive an audio command and generate at least one terminal recognized audio command, wherein the at least one terminal recognized audio command has a corresponding terminal confidence value;

5

at least one network speech recognition engine operably coupled to the

       microphone and coupled to receive the audio command and generate at

       least one network recognized audio command, independent of the terminal

       speech recognition engine, wherein the at least one network recognized

       audio command has a corresponding network confidence value;

a comparator operably coupled to the terminal speech recognition engine operably

       coupled to receive the at least one terminal recognized audio command

       and further operably coupled to the at least one network speech

       recognition engine operably coupled to receive the at least one network

       recognized audio command; and

a dialog manager operably coupled to the comparator, wherein the comparator

       selects at least one recognized audio command having a recognized

       confidence value from the at least one terminal recognized audio

       command and the at least one network recognized audio command based

       on the at least one terminal confidence value and the at least one network

       confidence value, wherein the selected at least one recognized audio

       command is provided to the dialog manager;

a dialog manager audio command determined by the dialog manager from the at

       least one recognized audio commands based on the at least one recognized

       audio command confidence levels such that the dialog manager executes

       an operation in response to the dialog manager audio command; and

the dialog manager being operably coupleable to a content server such that the

       operation executed by the dialog manager includes accessing the content

       server.

6

RESPONSE

Applicant respectfully traverses and requests reconsideration.

Applicant's attorney wishes to extend gratitude to the Examiners for courtesies extended in the telephone interview conducted on October 22, 2002.

Applicant respectfully submits, for the Examiner's consideration, amended claims 2, 4-6, 8, 13, 15, 18, and 21. Applicant also presents for consideration new claims 23-26. Claim 23 is presented for examination, wherein claim 23 includes originally presented claim 1, including the limitation originally presented in claim 2, and further providing claim subject matter to the access of a content server. Claim 24 represents originally presented claim 7, including further patentable subject matter, claim 25 represents originally presented claim 12, including further patentable subject matter, and finally claim 26 includes originally presented claim 19 and further patentable subject matter.

It is respectfully submitted that these amendments do not present any new subject matter, and provide for the claimed limitation of inherently contained features of originally presented claims. As such, Applicant respectfully submits that the amendments are not narrowing in nature, but merely a further delineation of inherently contained features already therein. Should the Examiner feel that this amendment is narrowing in nature, Applicant respectfully requests an express assertion of the Examiner's position.

Regarding added claim 23, the limitation of "selecting at least one recognized audio command" was originally presented in claim 2 and the limitation of "accessing a content server" is presented within the specification at, among other places, page 14, lines 16-19. As such, Applicant once again respectfully submits that the amended claims do not add any new subject matter and are fully supported by the specification, as filed.

Claims 1, 2, 4-6 and 12-18 originally currently stood rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent Application Publication 2002/0091518A1 having inventors Baruch, et al. (hereinafter referred to as "Baruch") in view of U.S. Patent No. 6,006,183 issued to Lai, et al. (hereinafter referred to "Lai").

7

Claims 1 and 12 (currently pending claims 23 and 24) were directed to, among other things, a method and apparatus for multi-level distributed speech recognition. The method and apparatus utilizes a first speech recognition engine and an at least one second speech recognition engine, wherein both speech recognition engines recognize an audio command and thereupon generate at least one recognized audio command output. The method and apparatus further includes, *inter alia*, selecting at least one of the recognized audio commands based on an associated confidence value and accessing a content server in response to the recognized audio command.

Baruch teaches, among other things, a control unit that includes a recognition result receiver capable of receiving recognition results, a recognition result association unit and a recognition engine activator capable of activating the recognition engine associated with a recognition result. Baruch teaches, *inter alia*, having a plurality of varying types of speech recognition engines, wherein based on a specific type of input, at least one of the specific types of engines is activated for the purpose of recognition. For example, Baruch discloses that upon the recognition of the command DIAL by a SICC engine 26, the system 10 may switch into a digit dialing mode, which may include the SIDD-M and/or the SDDD-IL modes. (¶ 43) Therefore, Baruch teaches a system that selects between multiple speech recognition engines based upon an anticipated type of speech input, wherein upon final speech recognition, the recognized speech command is provided either to a digital communication unit 30 or display unit 32. Moreover, "if the voice input is recognized as one of the target names, digital communication unit 30 may perform the action and may be required to establish connection with the target, for example, by dialing a telephone number." (¶ 40)

Lai discloses a speech recognition confidence level display system wherein a user provide a speech input to a microphone 170 which is thereupon provided to the speech engine 160. The speech engine produces a plurality of words/scores, 220, 230 which are provided to a graphical user interface application 150 having a confidence level indicator process 180. In response to a user control 140 within a graphic user interface display 105, the GUI application 150 provides to the display multiple words having associated

8

attributes (110, 120 and 130) for the user to provide a visual interface of the associated recognition confidence value.

It is respectfully submitted that the combination of Baruch in view of Lai, fails to teach or suggest all of the claimed limitations of the present claimed invention. Among other things, Baruch fails to disclose "accessing a content server in response to the at least one recognized audio command." As stated above, Baruch teaches, at best, "dialing a telephone number." (¶ 40). The present invention clearly discloses the claimed limitation of accessing a content server, wherein the combination of Baruch and Lai fails to disclose accessing a content server and provide for, at most, dialing a telephone number with the digital communication unit 30 to access a specific person, or providing an output display unit 32, which is inconsistent with the claimed limitation of, among other things, accessing a content server in response to the at least one recognized audio command." Furthermore, it is respectfully submitted that Baruch teaches, selecting a selected speech recognition engine, in response to a first recognized audio command, which is inconsistent with the claimed present invention of "accessing a content server."

In the present Office Action, on page 7, the Examiner asserts that regarding claim 15, "Baruch teaches that through voice commands a user can access a list of previously selected languages where the list may be provided over a loud speaker" (¶ 44), which corresponds to "wherein the dialog manager accesses a content server and retrieves encoded information in response to the dialog manager audio command." Applicant respectfully traverses the Examiner's assertions made herein and must respectfully disagree. It is respectfully submitted that the Examiner-cited passage is inconsistent with the claimed limitation because, *inter alia*, Baruch teaches allowing the user to choose from a list of possible languages in a set-up mode. The only teaching Baruch provides in the Examiner-cited passage consists of verbal navigation commands (e.g. UP or DOWN) or allowing a user to select a language based on the speaking the name of the language. The examiner-cited passage fails to teach or suggest, *inter alia*, accessing the list of possible languages in response to an audio command and further fails to teach or suggest, *inter alia*, the access of a content server to retrieve this language information.

9

Regarding claims 2, 4-6, 13 and 15-18, Applicant respectfully submits that these claims contain further patentable subject matter in view of the combination of Baruch and Lai. For example, claim 2 recites "receiving encoded information from the content server." As stated above, Baruch and/or Lai fail to teach or suggest, among other things, a content server, therefore the structure and accompanying limitations regarding receiving encoded information from the content server are not disclosed. As such, Applicant respectfully submits that claims 2, 4-6 and 13-18 contain patentable subject matter in and of themselves, in view of the teachings of Baruch in view of Lai.

Regarding claims 1, 12, and 14, Applicant respectfully submits that the present rejection is no longer applicable, as claims 1, 12, and 14 have herein been cancelled, without prejudice.

Applicant respectfully requests reconsideration and withdrawal of the present rejection of claims 2, 4-6, 13, 15-18 and 23-24. Furthermore, Applicant respectfully requests the passage of these claims to issuance.

Claims 3, 7-11 and 19-22 stood rejected under 35 U.S.C. 103(a) as being unpatentable over Baruch in view of Lai and further in view of U.S. Patent No. 6,122,613 issued to Baker (hereinafter referred to "Baker"). Applicant respectfully traverses and requests reconsideration.

Baker discloses, *inter alia*, the use of multiple speech recognizers being selectively applied to the same input sample. More specifically, Baker utilizes a real time recognizer 33 for high speed, but error-laden, voice recognition and an off line recognizer 309 for low speed, error-free transcription. The system utilizes a combiner 311 to generate a more accurate speech recognition result, the system further includes the ability to use an off line transcription station 313, such as an individual, which provides for further error free transcription. Regardless thereof, in response to the recognition results of the real time recognizer 303 and the off line recognizer 309, the combiner 311 with or without the off line transcription station 313, merely provides a stated output back to a monitor 305 for display to the user of the speech input. Generally speaking, Baker teaches a speech to text recognition system wherein automatic transcription may be

10

provided to a display, such as a video monitor 305. Baker fails to disclose, among other things, performing any further functions beyond the speech recognition, than merely providing the output to the display for the user's benefit and/or correction abilities to produce a final written document.

Regarding claim 3, Applicant respectfully resubmits the above position offered with regards to claims 23 and 2, and further submit that claim 3 contains further patentable subject matter therein.

Regarding claims 7 and 19 (currently pending 25 and 26), Applicant respectfully traverses the Examiner's characterization and the application of the prior art references with regards to the claimed limitations. Among other things, the combination of Baruch, Lai and Baker fail to teach or suggest the limitation of "accessing a content server in response to the at least one recognized audio command." As discussed above with regards to claims 23 and 24, Applicant respectfully resubmits that Baruch teaches a system that, upon recognition, either activates another speech recognition engine, provides an output to a display unit, or dials a telephone number with the digital communication unit 30, Lai, upon speech recognition, provides the multiple outputs with their accompanying attributes 110, 120 and 130 to a GUI display 105, and Baker, based upon the speech recognition, provides a visual output on a monitor 305. As such, none of the references, either individually or in combination thereof, teaches or suggests all of the claimed limitations.

Regarding claims 8-11 and 21-22, it is respectfully submitted that these claims contain further patentable subject matter in view of the combination of Baruch, Baker and Lai. For example, claim 8 recites "prior to accessing a content server, generating an error notification." As discussed above, none of the prior art references discloses accessing a content server, therefore these claims contain further patentable subject matter in view of the prior art of record.

Regarding claims 7 and 19-20, Applicant respectfully submits that the present rejection is no longer applicable, as claims 7, and 19-20 have herein been cancelled, without prejudice.

11

As such, Applicant respectfully requests reconsideration and withdrawal of the rejection regarding claims 3, 8-11, 21-22, and 25-26. Furthermore, it is respectfully requested that these claims be passed to issuance.

Attached hereto is a marked-up version of the changes made to the claims by the current amendment. The attached page is captioned "Version with markings to show changes made."

Accordingly, Applicant respectfully submits that the claims are in condition for allowance and that a timely Notice of Allowance be issued in this case. The Examiner is invited to contact the below-listed attorney if the Examiner believes that a telephone conference will advance the prosecution of this application.

Respectfully submitted,

By: _____
Timothy J. Bechen
Registration No. 48,126

Date: November 20, 2002
VEDDER, PRICE, KAUFMAN & KAMMHOLZ
222 N. LaSalle Street
Chicago, IL 60601
(312) 609-7500
FAX: (312) 609-5005